Commissioned by

The Wallace Foundation®
Supporting ideas.
Sharing solutions.
Expanding opportunities.®

ctp
Center for the Study of Teaching and Policy
UNIVERSITY OF WASHINGTON

Improving Leadership for Learning ■ **Stories from the Field**

Making State Accountability Count

How New Mexico supports principals with data tools

By Ron Feemster

Ron Feemster is a freelance writer who teaches journalism at Brooklyn College. He can be reached at ronfeemster@gmail.com.

Linda Paul, the school district superintendent in Aztec, NM, was unhappy when she received her students' test scores from the Public Education Department in 2005, but not because too few of her students had passed the state's assessment exam. Paul was frustrated, as she had been for several years, because the data were nearly incomprehensible.

New Mexico's standards-based assessment test (NMSBA) determines which schools are making the adequate yearly progress (AYP) required to comply with the federal No Child Left Behind (NCLB) law. Each year, the state raises the percentage of students who must be proficient in math and reading to achieve AYP. Districts that fail to make AYP are designated in need of improvement. If they fail to measure up for several years in a row, schools come under more scrutiny and, eventually, under direct control of the state. The state measures percentage improvement in each subject area separately for eight groups of students: Hispanic, Native American, white, African-American, Asian, English language learners, students with disabilities (including special education students), and low-income students. Elementary schools also can fall short in attendance, and high schools may fail based on graduation rates. The AYP target percentages increase each year until 2013–14, when every student must be proficient.

So for Paul, like all other superintendents, the stakes are high when the state issues test scores. However, all she received from the Public Education Department in Santa Fe was a big box of paper printouts and a CD crammed with raw data. The printouts contained a data line for each of the 3,500 students in her sprawling district among the natural gas fields and Indian reservations on the Colorado border. But she found no summary data for her six schools — no tabulations by grade level, subject matter, sex, or ethnicity. The information she needed to help her principals plan instruction was missing. She was able to open the data CD, but all she saw was a text file full of letters, numbers, and commas — no manual, no "readme" files, not even a slip of paper to suggest what program might help her make sense of it all.

"The data were useless," Paul remembers. She struggled to understand the legend used to decode scores. The data line for little Jose, a 4th grader, might state that he got a 30 on benchmark one. But it was not obvious what the benchmark measured. Nor could Paul determine if 30 was a high or low score. She was unable to diagnose the problems in her schools.

"We needed to be able to ask questions of the data," she says. "How are my 4th grade Hispanic students doing in reading? When they have specific problems, what are they? We needed answers to help us plan."

Happily, help arrived from Santa Fe the same day that box of useless paper printouts landed on Paul's desk. Even as Paul sat poring over the data and sorting students by hand into grade levels and schools, Peter Winograd, the director of the state's Education Accountability Office, dropped in to discuss a new program to increase "data literacy" among education leaders. Funded by The Wallace Foundation, the program aimed to transform the vast amounts of data collected during assessment tests into a tool that leaders could use to set policy and, above all, shape instruction.

"We are not sitting around in Santa Fe to try and 'catch' schools that are not doing well," Winograd explains. "Our goal is to help education leaders lead. We go around the state asking how we can help."

"What you want to do is make sure that principals, superintendents, and other education leaders can use the data in the warehouse effectively to improve student achievement."

New Mexico school districts need a lot of help. One in five of New Mexico's 328,000 students is learning English as a foreign language, according to *Journey to Excellence*, a report produced this year by Winograd's office and the New Mexico Business Roundtable for Educational Excellence. One in five has a disability, and one in four students lives in poverty, which is the greatest barrier to student achievement, according to the report. In some districts, particularly in the southern part of the state, many students have at least one parent who entered the United States illegally. Nearly half of New Mexico's students live in five of the state's 89 school districts. More than half of the state's 801 schools are located in rural areas, often with out-of-date facilities, long bus routes, and few selling points when it comes to recruiting teachers. In the 2006–07 school year, 429 schools (53 percent) failed to meet AYP for NCLB. In 2004–05, when NCLB set the AYP bar lower, just 32 percent of New Mexico's schools failed to clear the hurdle. The rapidly rising AYP standard is increasing the pressure on school administrators to improve student achievement. The state hopes that easier access to better data will help schools reverse the downward trend and keep pace with federal requirements.

Paul gave her mysterious data CD to Winograd and his senior policy analyst, Beata Thorstensen, who in turn found a university professor to help unpack the data. The CD turned out to be comma-delimited ASCII data, a format commonly imported into Excel, a Microsoft spreadsheet program that almost all school districts own. The questions that Paul and many other educators wanted to ask could be posed and answered with pivot tables, an Excel function that extracts specific reports from raw data in Excel.

"We were delighted when we discovered pivot tables," Winograd says. "We did not have to ask districts to buy any proprietary software. They already had what they needed."

What some districts did not have was an expert to build the pivot tables. Winograd saw to it that superintendents in the Wallace project's six demonstration districts formed relationships with people who understood their questions and helped them query the data. The result

has been a simple point-and-click set of tables that can break down student performance not just by grade and subject matter, but by particular teachers, individual students, even by the individual topics on the test.

These topics, known as "benchmarks," are the smallest unit of testable knowledge on the NMSBA for which New Mexico releases student scores. A benchmark may be tested with a half dozen questions or more. Unlike neighboring Texas, the state does not publish the actual questions on the test. And while many New Mexico educators disdain what they call the "drill and kill" philosophy of Texas schools that teach to the test, most spend a great deal of time puzzling out the types of questions likely to surface in vaguely worded topics, such as these for 5th graders: "analyze cause and effect relationships" and "distinguish between fact and opinion."

Even when the exact content of the exam is a bit uncertain, charting student performance against benchmarks is a wonderful tool. Paul, like most other superintendents in the state, would love to use state test results for planning. But so far, the data have not been delivered in time. The NMSBA tests are administered every March. Because students must write out answers to many questions in both math and language arts, grading the test takes time. Although the state is getting faster, the results have yet to be published before school starts in August. Principals cannot yet exploit test data to group students by strengths or weaknesses.

Using a short-cycle assessment

"Schools are running from March to March, from yearly test to yearly test," Paul says. "And the cycle is just too long." Instead of trying to second-guess the test results and plan with too little data, Paul asked Princeton Review to create a short-cycle assessment test that the district administers to all students at the beginning, middle, and end of the school year. The short cycle results typically come back within two weeks. At $30,000 to $40,000 a year, the tests cost about $10 per student per annum. So far, the investment has paid off: Student performance on the second short-cycle test, given in January, successfully predicted test results on the NMSBA, given in March, which gave Paul confidence that it accurately identifies students who need help.

"Schools have two months to intervene and work on problem areas before the state test," Paul says. Like Aztec, most of the state's 89 districts are moving toward some kind of short-cycle assessment, Winograd says. But many cannot afford to hire a pricey vendor like Princeton Review.

This September, just two years after Winograd interrupted her attempt to correlate data by hand, Paul received her third set of pivot tables from NMSBA data. She also began a new round of short-cycle assessments.

"It's become a data-rich environment," she says. "I've always asked principals the same two questions: 'How's business?' and 'How do you know?' These days they all get out the data to show me." Before long, school superintendents and principals across the state will have access to better data through the Student Teacher Accountability Reporting System (STARS), a data warehouse the state is building. STARS will keep a running account of budget, student-achievement, and staff data.

"Data is the current currency," says Winograd, who has reached out to 140 principals in 30 districts to get feedback from end users of the new warehouse. "What you want to do is make sure that principals, superintendents, and other education leaders can use the data in the warehouse effectively to improve student achievement."

The new emphasis on data has changed the culture of accountability in the Aztec district. Just as the superintendent asks principals for their data, so the principals expect teachers to account for student performance. "Now," says Paul, "the facts are right there in the numbers. If a teacher is having trouble teaching algebra, we can see it. Data has helped us get egos out of teacher evaluations." What principals ask of teachers, teachers ask of students. As she showed a visitor through one of her elementary schools, Paul pointed to a three-ring binder on a 2nd grader's desk. My Data Book, the cover read.

"Everybody in the system is accountable for charting their own data," Paul says. "The students learn it very early."

From superintendents to 2nd graders, access to more meaningful data has transformed school districts throughout the state. In the Gadsden school district, which stretches from the border of Mexico and the city limits of El Paso, TX, halfway to Las Cruces — an area larger than Rhode Island — principals once ruled their schools like private fiefdoms.

"We needed to be able to ask questions of the data — How are my 4th grade Hispanic students doing in reading? ... We needed answers to help us plan."

"People used to make decisions by gut and by heart," says Yvonne Lozano, Gadsden's assistant superintendent for education. Teachers had a similar autonomy. They did things their own way and pointed to successful students as ones who "got it" or were particularly bright. When students did not progress, poor teaching and failing instructional programs were not the most likely suspects.

"Sometimes, teachers would just appeal to history," Lozano says. " 'We've always done it this way.' " The problem with what they had always done? It had not usually worked. In fact, the district's academic performance was ranked 88th of 89 districts in New Mexico in spring 1999. People made excuses: Many students were desperately poor. They came from the colonias, ghettos where people lacked not only steady work, but also running water, indoor plumbing, and often floors in their dwellings. When the parents lived on $5,000 a year in a colonia, how could a public school educate the child? The district also had more than its share of recent immigrants and English language learners.

Lozano and her colleagues in the central office worried that the schools had given up. They led a "self-takeover" of the district in the 1999–2000 school year. First, they instituted strong central control in the district, where 14,000 students enroll in three preschools, 14 elementary schools, three middle schools, and four high schools. Lozano wanted to "treat the district more like a business," so they took control of principals' funds and monitored their budgets. They began to hire better teachers and take steps to eliminate those who were not performing well. The district began to inch up the rankings.

As NCLB began to generate more data, the district tried to use it. "Looking harder at data was part of our attempt to behave more like a business," Lozano says. But it was not until it joined Winograd's program in 2005 that the district made a breakthrough. "The Wallace project allowed us to look deeper into the data provided by the state," Lozano says.

Once again, the key to exploiting data was creating pivot tables. The district asked Ken Korn, a retired teacher from nearby El Paso, to create pivot tables. The results mirrored Aztec: Administra-

tors quickly targeted specific areas in which students scored too low. Principals identified teachers who needed help and sent coaches into the classroom to model new lesson plans and share tips.

In Gadsden's system, the assistant principals' job description includes assisting in accountability. Lozano sent all 22 assistant principals in the system to Las Cruces, where Thorstensen trained them to use pivot tables. When they returned and brought their questions to Korn, he refined the tables to answer their new data queries. Before long, Lozano was calling the assistant principals "Ken Korn clones." The decision to focus on assistant principles had another important consequence: Gadsden had increased the data literacy of future principles in the district.

By 2006, the assistant principals had dug deep into the state testing data. For each of their 14,000 students, the district now had a benchmark profile. Every student's score on each benchmark was mapped not only against the total points available in that benchmark, but also against the score of the average proficient student. The results often were surprising.

"Many students who were not proficient in math or reading actually scored higher than the average proficient student on one or more benchmarks," Korn says. "We learned that students who weren't good enough on average were often very good at a few things. They weren't unable to learn. They already excelled at something. There were just a lot of things they hadn't been taught."

The next step will be to create a historical record for each child with respect to specific benchmarks. Even now, in the third year of NMSBA pivot tables, it has become possible to do some longitudinal work. Coming up with the right question may be the challenge. Korn is convinced the district has the data to answer it.

Charting progress of every student

This fall, some of the assistant principals trained in 2005 took over schools of their own. One of those new principals is Linda Perez, who took over Sunland Park Elementary School. As she had done as an assitant principal, she immediately set up a kind of "war room" to chart the progress of every student against key reading benchmarks. With only a few hundred students in her school, this chart takes up an entire wall. She and every teacher in the building can tell at a glance where any student stands.

Perez, the data maven, has other worries at her school, however. Before she showed a visitor the progress board charting student achievement, she pointed through the rear door of her cheerful turquoise school building to the chain-link fence marking the Mexican border. A border patrol lookout post stands atop a nearby bluff. A few times, she said, a man or woman has sprinted past her school on the way to America, presumably without a visa.

"The border patrol sometimes stops parents on the way to pick up their children," Perez notes coolly. "It makes it harder to get parental involvement at school."

What draws parents in, whatever their fears or misgivings, is student success. And Gadsden has made great strides in the last three years. In the latest ranking, Gadsden rose to third among 89 districts, just behind Albuquerque and Las Cruces. The businesslike, data-driven district is succeeding. "If you have a data sheet on each child year after year," Korn says, "you have a commitment to leaving no child behind."

This article is one of a series written by education journalists highlighting useful practices in improving education leadership. The series, part of a research project conducted by the Center for the Study of Teaching and Policy at The University of Washington, is supported by a grant from The Wallace Foundation. Opinions represent those fof the authors and not necessarily those of the foundation.